INTERNATIONAL STANDARD

ISO/IEC 19566-6

First edition
2019-07

Information technologies — JPEG systems —

Part 6: JPEG 360

Technologies de l'information — Systèmes JPEG JPEG 360 —

Partie 6: JPEG 360

Reference number
ISO/IEC 19566-6:2019(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 19566 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The ISO/IEC 19566 series is designed primarily for format and metadata storage and protection method of compressed continuous-tone photographic content.

There is increasing use of multi-sensor images from multiple image sensor devices, such as 360 degree capturing cameras or dual-camera smartphones available to consumers. Images from these cameras are shown on computers, smartphones, and head-mounted displays (HMD).

Because existing JPEG standards do not fully cover these new uses, incompatibilities have reduced the interoperability of these images, and thus reducing the widespread ubiquity which consumers have come to expect when using JPEG-based images.

Additionally, new modalities for interacting with images, such as computer-based augmentation, face-tagging, and object classification require support for metadata that was not part of the original JPEG scope.

This document defines "JPEG 360", building upon the features of JPEG Universal Metadata Box Format (JUMBF) (see ISO/IEC 19566-5) which itself builds upon ISO/IEC 18477-3 (Box file format) which provides compatibility with ISO/IEC 10918-5 (JPEG File Interchange Format (JFIF)).

This document defines the use of the JPEG 360 Content Type JUMBF superbox with respect to the sub-box components which include the definition of an XML box, the use of other boxes such as unstitched image elements for omnidirectional captures together with the main image and descriptive metadata, and encrypted parts of the image.

Information technologies — JPEG systems —

Part 6:
JPEG 360

1 Scope

This document specifies omnidirectional/360-degree image and motion contents using Rec. ITU-T T.81 | ISO/IEC 10918-1, Rec. ITU-T T.800 (11/2015) | ISO/IEC 15444-1, and ISO/IEC 18477-3.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 19566-5, *Information technology — JPEG Systems — Part 5: JPEG Universal Metadata Box Format (JUMBF)*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1.1
box
binary structure that encapsulates an object embedded in a file

3.1.2
box-based file format
file format whose composing elements are boxes containing structured data in compliance with ISO-based media file format

3.1.3
deserialization
extraction of data structure from a series of bytes

3.1.4
equirectangular projection
projection for mapping a portion of the surface of a sphere to a flat image

3.1.6
metadata
data that describes other data, including text, image, hypertext and combinations thereof

3.1.7
omnidirectional
(sub)spherical surface of an image of a scene as if observed from a single point of projection.

3.1.8
serialization
translation of data structures into a series of bytes that can be stored and/or transmitted

3.2 Abbreviated terms

ASCII	American Standard Code for Information Interchange
DCT	discrete cosine transform
ERP	equirectangular projection
FOV	field of view
JSON	JavaScript object notation
JPEG	joint photographic experts group
JUMBF	JPEG universal metadata box format
MIME	multipurpose internet mail extensions
URL	uniform resource locator
XML	eXtensible Markup Language
XMP	eXtensible Metadata Platform
RDF	resource description framework
W3C	World Wide Web Consortium
umf	universal metadata framework

4 Conventions

4.1 Conformance language

In this document, the following verbal forms are used:

— “shall” indicates a requirement;

— “should” indicates a recommendation;

— “may” indicates a permission;

— “can” indicates a possibility or a capability.

Information marked as “NOTE” is intended to assist the understanding or use of the document. “Notes to entry” used in Clause 3 provide additional information that supplements the terminological data and can contain provisions relating to the use of a term.

The keyword "reserved" indicates a provision that is not specified at this time, shall not be used, and may be specified in the future. The keyword "forbidden" indicates "reserved" and in addition indicates that the provision will never be specified in the future.

4.2 Typesetting

Regular face fonts as this text describe informative text that provides instructions, comments or details for the reader.

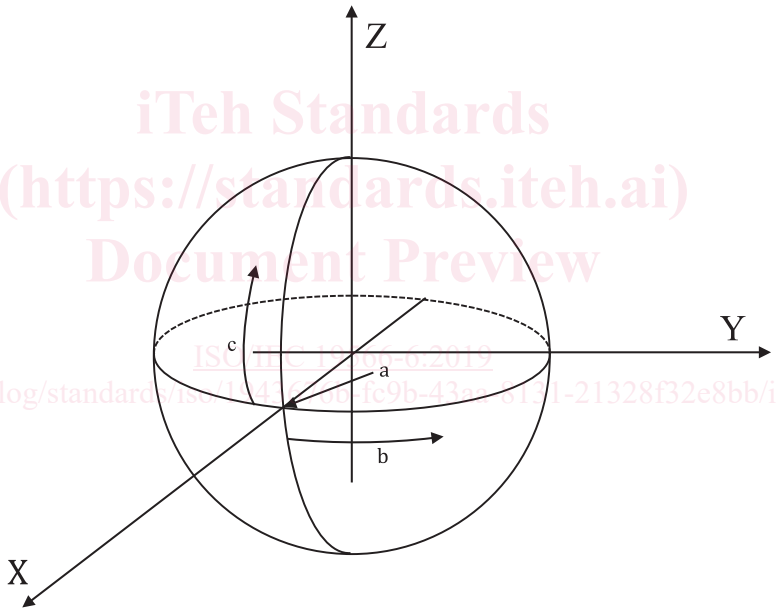
```
Monospaced text as this paragraph indicates program input or output as necessary to either
run the software, or as generated by the software on the console.
```

NOTE The character values of the monospaced text in this document could differ the actual value due to differences in the character encodings used; e.g., IS0 8859 vs UTF-8.

5 Description and definition of JPEG 360 images

The equirectangular projection is a commonly used projection of omnidirectional cameras to a two-dimensional rectangular image and is the default image projection for JPEG 360.

The equirectangular projection maps the image onto a spherical surface from a single projection point at the centre of the sphere. The surface of the sphere is described by two angular measures as shown in Figure 1; for convenience, a unit sphere is assumed. The centre of the sphere is coincident with the origin of the three-dimensional Cartesian coordinates. A point on the surface of the sphere is defined by the two angles ϕ and θ, which are also referred to as “longitude” and “latitude” respectively.



Key

a origin of 360 image (θ = ϕ = 0)

b direction of increasing ϕ

c direction of increasing θ

Figure 1 — Spherical and Cartesian coordinates

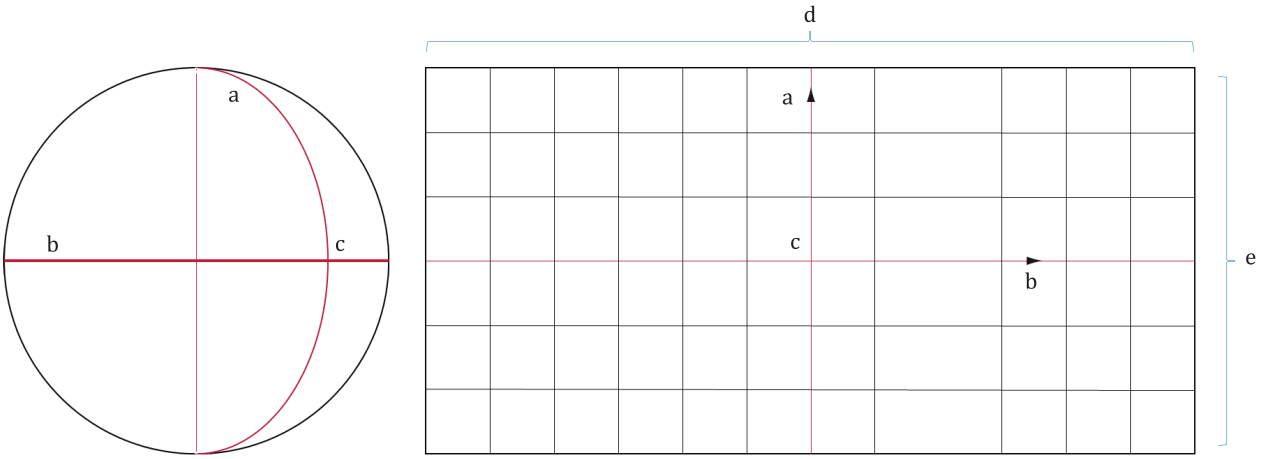
In Cartesian coordinates, the points on the sphere are defined as follows:

— $x = \cos\phi \cdot \cos\theta$

— $y = \sin\phi \cdot \sin\theta$

— $z = \sin\theta$

For an equirectangular projection the sphere's surface, expressed in terms of the angles ϕ and θ, can be mapped to a two-dimensional Cartesian surface. For example, in Figure 2 which includes the mapping for several lines of constant longitude or constant latitude.



Key

a prime meridian (commonly at ϕ = 0)
b equator (typically at θ = 0)
c origin of 360 image (θ = ϕ = 0)
d span of ϕ (commonly −180° to 180°)
e span of θ (commonly −90° to 90°)

Figure 2 — Description of mapping from spherical surface to equirectangular

For this document, a simplified representation will be used, for example, for equirectangular projection for a full spherical surface as shown in Figure 3.

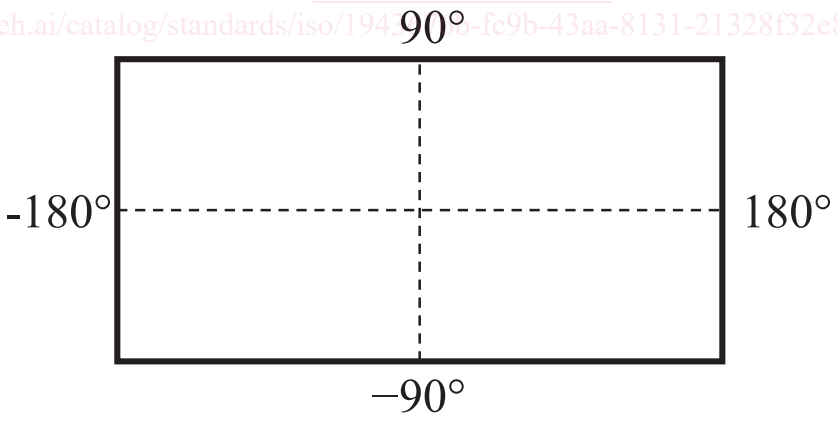


Figure 3 — Simplified representation for equirectangular projection for full spherical surface

The generalized equirectangular projection is expressed in terms of ranges for ϕ and θ by defining four values ϕ_{min}, ϕ_{max}, θ_{min}, and θ_{max} as shown in Figure 4.

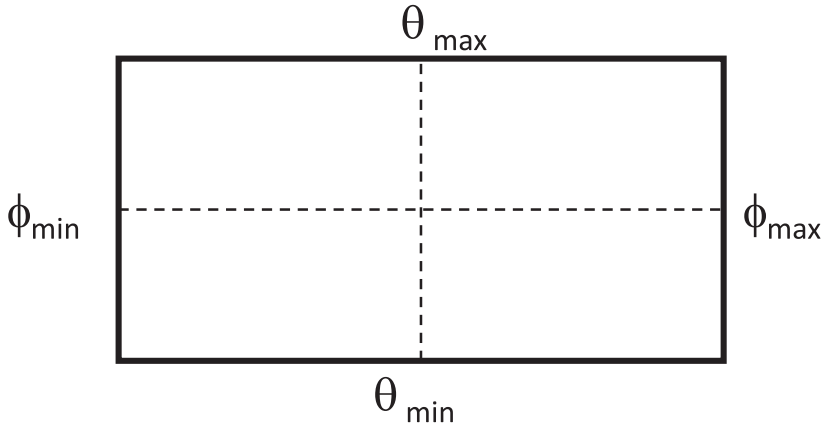


Figure 4 — Generalized description of equirectangular projection

ϕ_{min}, and ϕ_{max} are constrained by the two conditions:

— $-360 \leq \phi_{min} \leq \phi_{max} \leq 360$, and

— $\phi_{max} - \phi_{min} \leq 360$

θ_{min}, and θ_{max} are constrained by the two conditions:

— $-180 \leq \theta_{min} \leq \theta_{max} \leq 180$, and

— $\theta_{max} - \theta_{min} \leq 180$

The advantages of this generalized description are:

i) it defines equirectangular projections that are smaller than a full sphere, and

ii) it can mathematically shift the origin (0, 0) of the equirectangular project in terms of the angle range values.

For example, an “equatorial image” to be a band around the equator of the equirectangular projection, or set the origin at the left edge of the equirectangular projection.

5.1 Relative orientation

The full 360 camera captures equivalent scenes regardless of the exact position the user holds the camera during image exposure. However, it is highly desirable to improve viewer’s experience of the image by displaying the image in a way that is nominally oriented so that the scene horizon is parallel to the equirectangular projection’s centre line. Sensors in the camera can provide the camera orientation relative to the local direction of the Earth’s gravity; using this information allows the image to be remapped to the expected view.

In Figure 5 shows the direction of Earth’s gravity with respect to the camera’s local coordinate systems in both 3D space and equirectangular projection for a full spherical surface as might be expected when the camera is held in its upright position.